Memory technology evolution: an overview of system memory technologies

Technology brief, 9th edition

Introduction

This paper gives you an overview of the memory technologies that we use in HP ProLiant servers, and describes how we evaluate these technologies. It also briefly summarizes the evolution of server memory and explores the different dynamic random access memory (DRAM) technologies.

Processors use system memory to store the operating system, applications, and data they use and manipulate. As a result, speed and bandwidth of the system memory controls application performance. Over the years, the need for greater memory bandwidth has driven system memory evolution from asynchronous DRAM technologies to high-bandwidth synchronous DRAM (SDRAM), and finally to today's Double Data Rate (DDR) SDRAM technologies. Our challenge going forward is to continue to increase system performance by narrowing the performance gap between processors and memory.

The processor-memory performance gap occurs when the processor idles while it waits for data from system memory. In an effort to bridge this gap, HP and the industry are developing new memory technologies. We work with the Joint Electronic Device Engineering Council (JEDEC), memory vendors, and chipset developers. We evaluate developing memory technologies in terms of price, performance, reliability, and backward compatibility, and then implement the most promising technologies in ProLiant servers.

Basic DRAM operation

Before a computer can perform any useful task, it must copy applications and data from the disk drive to the system memory. Computers use two types of system memory:

- Cache memory, which consists of very fast static RAM (SRAM) integrated with the processor.
- Main memory, which consists of DRAM chips on DIMMs packaged in various ways depending on system form factor.

Each DRAM chip contains millions of memory locations, or cells, arranged in a matrix of rows and columns (Figure 1). Peripheral circuitry on the DIMM reads, amplifies, and transfers the data from the memory cells to the memory bus. Each DRAM row, called a page, consists of several DRAM cells. Each DRAM cell contains a capacitor capable of storing an electrical charge for a short time. A charged cell represents a "1" data bit, and an uncharged cell represents a "0" data bit. To maintain the validity of the data, the DIMM recharges, or refreshes, the capacitors thousands of times per second.

Figure 1. Representation of a single DRAM chip on a DIMM
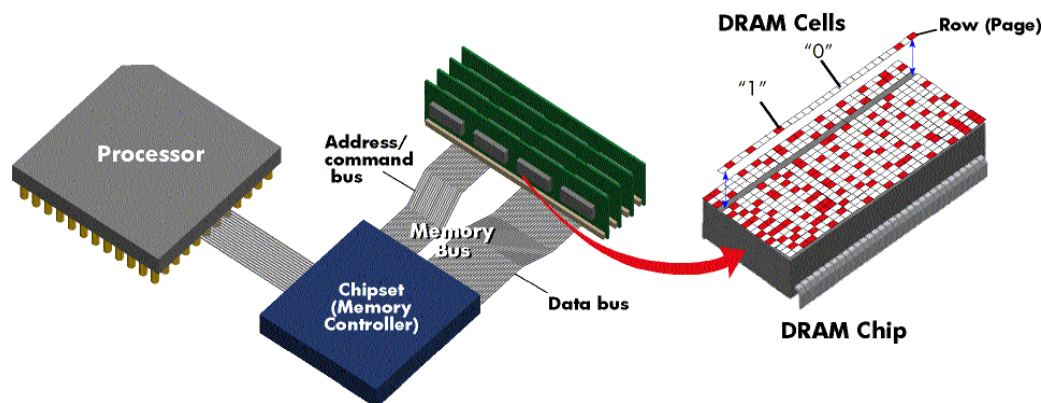
The memory subsystem operates at the memory bus speed. When the memory controller accesses a DRAM cell, it sends electronic address signals that specify the target cell's row address and column address. The memory controller sends these signals to the DRAM chip through the memory bus, which consists of:

- The address/command bus
- The data bus

The data bus is a set of lines, also known as traces, that carry the data to and from the DRAM chip. Each trace carries one data bit at a time. The throughput, or bandwidth, of the data bus depends on its width in bits and its frequency. The data width of a memory bus is usually 64-bits, which means that the bus has 64 traces, each of which transports one bit at a time. Each 64-bit unit of data is a data word.

The address portion of the address/command bus is a set of traces that carry signals identifying the location of data in memory. The command portion of the address/command bus conveys instructions such as read, write, or refresh.

When DRAM memory writes data to a cell, the memory controller selects the data's location. The memory controller first selects the page by strobing the Row Address onto the address/command bus. The memory controller then picks out the exact location by strobing the Column Address onto the address/command bus (Figure 2). These actions are Row Address Strobe (RAS) and Column Address Strobe (CAS). The Write Enable (WE) signal activates at the same time as the CAS to order a write operation. The memory controller then moves the data onto the memory bus. The DRAM devices capture the data and store it in the cells.

Figure 2. Representation of a write operation for FPM or EDO RAM
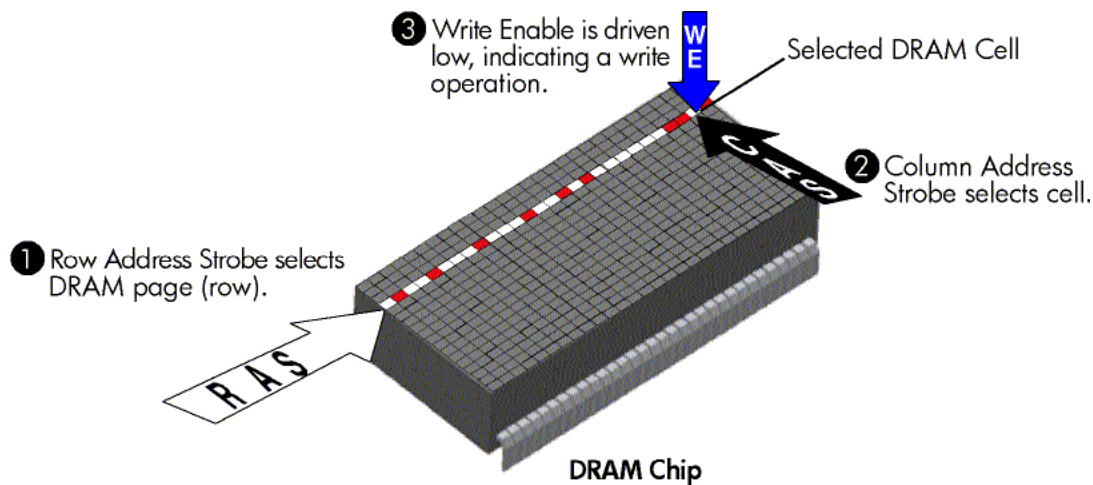
During a DRAM read operation, the memory controller drives RAS, followed by CAS, onto the memory bus. The WE signal is held inactive, indicating a read operation. After a delay called CAS Latency, the DRAM devices move the data onto the memory bus.

The memory controller cannot access DRAM during a refresh. If the processor makes a data request during a DRAM refresh, the data will not be available until the refresh completes. There are many mechanisms to refresh DRAM:

- RAS only refresh
- CAS before RAS (CBR) refresh, which involves driving CAS active before driving RAS active. CBR is used most often
- Hidden refresh

DRAM storage density and power consumption

DRAM storage capacity is inversely proportional to the cell geometry. In other words, the storage density increases as the cell geometry shrinks. Over the past few years, capacity has expanded from 1 kilobit (Kb) per chip to 2 gigabit (Gb) per chip. We expect that the capacity will soon grow to 4 Gb per chip.

The industry-standard operating voltage for computer memory components was originally 5 volts. But as cell geometries decreased, memory circuitry became smaller and more sensitive. Likewise, the industry-standard operating voltage decreased. Today, computer memory components operate at 1.8 volts, letting them run faster and consume less power.

Memory access time

The elapsed time from the assertion of the CAS signal until the data is available on the data bus is the memory access time or CAS latency. For asynchronous DRAM, we measure memory access time in nanoseconds. For synchronous DRAM, we measure memory access time by the number of memory bus clocks.

System bus timing

A system bus clock controls all computer components that execute instructions or transfer data. The system chipset controls the speed, or frequency, of the system bus clock. The system chipset also regulates the traffic between the processor, main memory, PCI bus, and other peripheral buses.

The bus clock is an electronic signal that alternates between two voltages (designated as "0" and "1" in Figure 3) at a specific frequency, measured in millions of cycles per second or megahertz (MHz). During each clock cycle, the voltage signal moves from "0" to "1" and back to "0." A complete clock cycle spans from one rising edge to the next rising edge. Data transfer along the memory bus can start on either the rising edge or the falling edge of the clock signal.

Figure 3. A bus clock signal
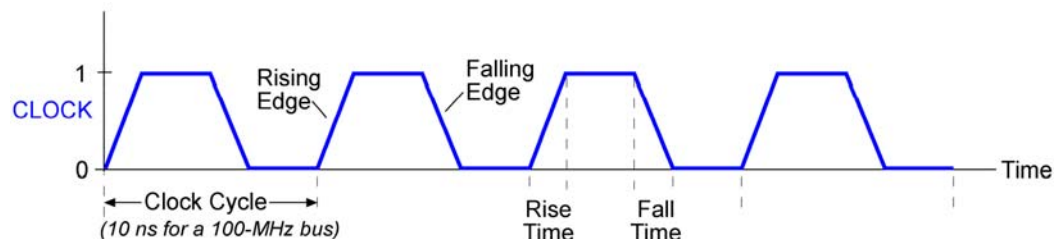
The clock cycle time is measured in nanoseconds (ns, 1×10^{-9} second) and the clock frequency is measured in megahertz (MHz, millions of cycles per second). The clock frequency is the reciprocal of the clock cycle time. For example, if the clock cycle is 10 ns then the clock frequency is 1/(0.000000010)=100 MHz.

System components run at different speeds from one another in a typical system. For this reason, different clocks running at various but related speeds control the components. These clocks use various clock multiplier and divider circuits to generate multiple signals. All these signals derive from the main system bus clock. For example, if the main system bus operates at 100 MHz, a divider circuit can generate a PCI bus frequency of 33 MHz (system clock ÷ 3) and a multiplier circuit can generate a processor frequency of 400 MHz (system clock x 4). Computer components that operate in whole multiples of the system clock are termed synchronous because they are "in sync" with the system clock.

Synchronous components operate more efficiently than asynchronous components. With asynchronous components, either the rest of the system or the component itself must wait one or more additional clock

cycles for data or instructions. In contrast, synchronized components know which clock cycle will make data available, eliminating these timing delays.

Memory bus speed

DRAM speed is not the same as the true speed, or frequency, of the memory subsystem. The memory subsystem operates at the memory bus speed, which may not be the same frequency (in MHz) as the main system bus clock. Memory timing and maximum DRAM speed control memory subsystem speed. The server BIOS sets the memory system speed based on the fastest common frequency of the loaded DIMMs.

Burst mode access

The original DRAM took about six system-bus clock cycles for each memory access. During memory access, first the RAS, CAS, and then 64 bits of data moved through the memory bus. The next sequential address access required a repeat of the RAS-CAS-Data sequence. As a result, most of the overhead occurred while transferring row and column addresses, not data.

Later DRAM implementations improved performance by retrieving data from sequential memory locations on the assumption that the system will use all of the data. Called burst mode access, this process accesses four consecutive 64-bit sections of memory, one after the other, based on the address of the first section. So instead of taking six clock cycles to access each of the last three 64-bit sections, it can take from one to three clock cycles each (see Figure 4).

Burst mode access timing is normally stated in the format "x-y-y-y" where "x" represents the number of clock cycles to read/write the first 64 bits and "y" represents the number of clock cycles required for the second, third, and fourth reads/writes. For example, prior to burst mode access, DRAM took up to 24 clock cycles (6-6-6-6) to access four 64-bit memory sections. With burst mode access, DRAM accesses three additional data sections with every clock cycle after the first access (6-1-1-1) before the memory controller has to send another CAS.

Figure 4. Burst mode access
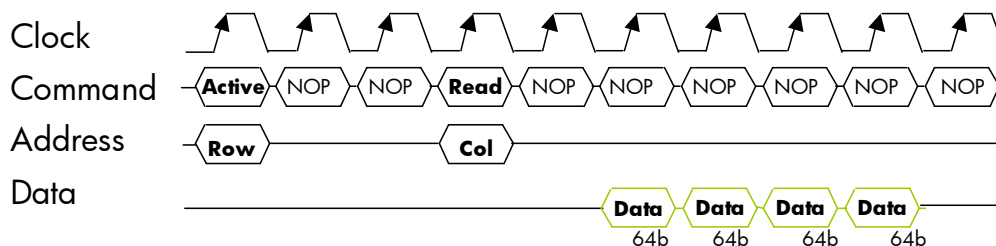
NOTE: NOP = No Operation Instruction

SDRAM technology

DRAMs are asynchronous, that is, without a memory bus clock. The memory controller determines when to assert signals and when to expect data based on absolute timing. Transferring data between a synchronous system bus and an asynchronous memory bus is inefficient and results in longer latency.

Consequently, the JEDEC developed the synchronous DRAM standard to reduce the number of system clock cycles required to read or write data. SDRAM uses a memory bus clock to synchronize the input and output signals on the memory chip. This simplifies the memory controller and reduces CPU-to-memory latency. In

addition to synchronous operation and burst mode access, SDRAM has other features that accelerate data retrieval and increase memory capacity:

- Multiple memory banks
- Greater bandwidth
- Register logic chips.

Bank interleaving

SDRAM divides memory into two to four banks for simultaneous access to more data. This division and simultaneous access is interleaving. Two-way interleaving is similar to dividing each page in a notebook into two parts and having two assistants retrieve a different part of the page. Even though each assistant must take a break, breaks are staggered so that at least one assistant is working at all times. Both assistants retrieve the data much faster than a single assistant does, especially since the single assistant does not access data when taking a break. This means that while the processor accesses one memory bank, the other bank stands ready for access. The processor can initiate a new memory access before the previous access completes, which results in continuous data flow.

Increased bandwidth

Memory bus bandwidth increases with its width (in bits) and its frequency (in MHz). By transferring 8 bytes (64 bits) at a time and running at 100 MHz, SDRAM increases memory bandwidth to 800 MB/s, 50% more than EDO DRAMs (533 MB/s at 66 MHz).

Registered SDRAM modules

To increase memory subsystem capacity, designers can embed register logic chips, referred to as registers, in DIMMs to act as pass-through buffers for address and command signals. With these registers, the memory controller does not have to drive the entire arrangement of DRAM chips on each module. Rather, the memory controller loads only the registers on each module. Each DIMM's register relays the address and command signals to the appropriate DRAM chip. Simultaneously, a phase-locked loop chip on a registered DIMM generates a second clock signal that runs synchronously with the system bus clock. This means the system bus clock does not have to drive all the DRAM chips. This lets you add more memory modules to the memory bus, increasing capacity.

DIMM configurations

DIMMs are available in several configurations, depending on the number and type of DRAMs used as well as the number of memory blocks, called ranks, the DIMM supports.

Single-sided and double-sided DIMMs

Each DRAM chip on a DIMM provides either 4 bits or 8 bits of a 64-bit data word. Chips that provide 4 bits are called x4 (by 4), and chips that provide 8 bits are called x8. Eight x8 chips or sixteen x4 chips make a 64-bit word, so at least eight chips are located on one or both sides of a DIMM. However, a standard DIMM has enough room to hold a ninth chip on each side. The ninth chip stores 4 bits or 8 bits of Error Correction Code (ECC).

An ECC DIMM with nine DRAM chips on one side is single-sided, and an ECC DIMM with nine DRAM chips on each side is double-sided (Figure 5). A single-sided x8 ECC DIMM and a double-sided x4 ECC DIMM each create a single block of 72 bits (64 bits plus 8 ECC bits). In both cases, a single chip-select signal from the memory controller activates all the chips on the DIMM. In contrast, a double-sided x8 DIMM (bottom illustration) requires two chip-select signals to access two 72-bit blocks on two sets of DRAM chips.

Single-rank, dual-rank, and quad-rank DIMMs

Along with single-sided and double-sided configurations, we classify DIMMs by rank. A memory rank is an area or block of 64-bits (72 bits for ECC memory) created by using some or all of the DRAM chips on a DIMM.

A single-rank ECC DIMM (x4 or x8) uses all of its DRAM chips to create a single block of 72 bits, and all the chips are activated by one chip-select (CS) signal from the memory controller (top two illustrations in Figure 5). A dual-rank ECC DIMM produces two 72-bit blocks from two sets of DRAM chips on the DIMM, requiring two chip-select signals. The chip-select signals are staggered so that both sets of DRAM chips do not contend for the memory bus at the same time. Quad-rank DIMMs with ECC produce four 72-bit blocks from four sets of DRAM chips on the DIMM, requiring four chip-select signals. Like dual-rank DIMMs, the memory controller staggers the chip-select signals.

Figure 5. Single-sided and double-sided DDR SDRAM DIMMs and corresponding DIMM rank
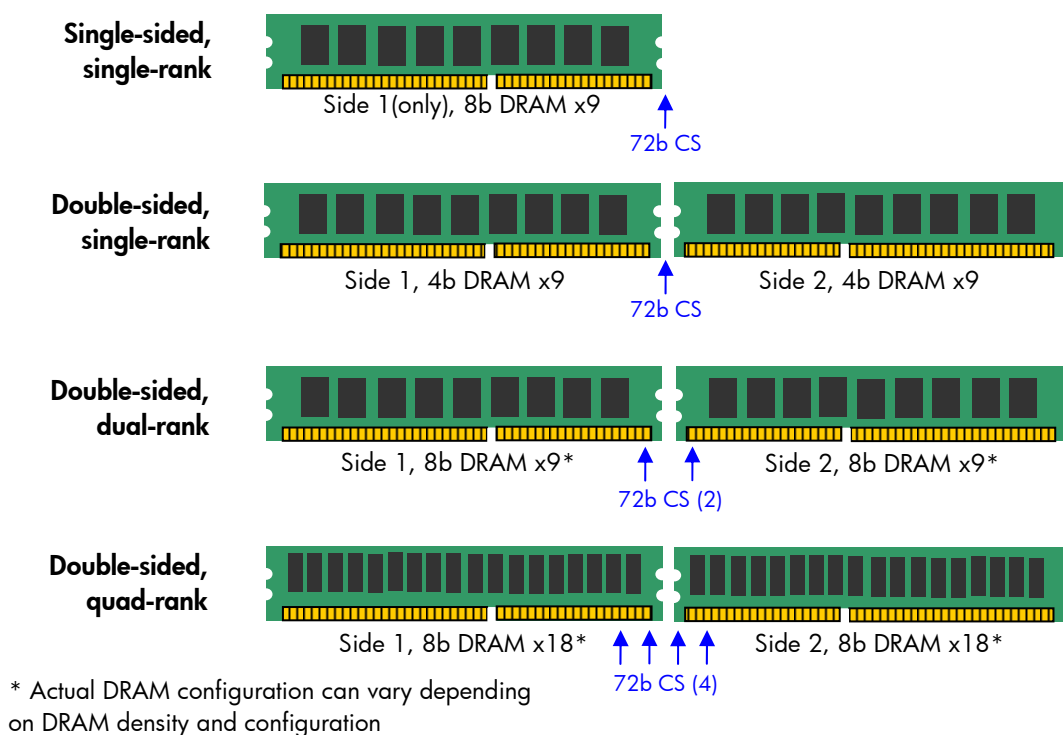
Memory ranks have become more important because of new chipset and memory technologies and larger memory capacities. Dual-rank DIMMs improve memory capacity by placing two single-rank DIMMs on one module. The chipset considers each rank as an electrical load on the memory bus. At slower bus speeds, the number of loads does not degrade bus signal integrity. For faster memory technologies, the chipset can drive only a certain number of ranks. For example, if a memory bus has four DIMM slots, the chipset may be capable of supporting only two dual-rank DIMMs or four single rank DIMMs. If you install two dual-rank DIMMs, then the last two slots must remain empty. To compensate for the reduction in the number of slots, chipsets now use multiple memory buses.

If the total number of ranks in the populated DIMM slots exceeds the maximum number of loads, the chipset can support, the server may not boot properly or it may not operate reliably. Some systems check the memory configuration while booting to detect invalid memory bus loading. If the system detects an invalid memory configuration, it stops the boot process to avoid unreliable operation.

To prevent such memory-related problems, we advise our customers to use only HP-certified DIMMs, which are available in the memory option kits for each ProLiant server (see the "Importance of using HP-certified memory modules in ProLiant servers" section).

Another important difference between single-rank and dual-rank DIMMs is cost. Memory costs generally increase with DRAM density. For example, the cost of an advanced, high-density DRAM chip usually runs more than twice that of a conventional DRAM chip. Because large capacity, single-rank DIMMs are manufactured with higher-density DRAM chips, they typically cost more than dual-rank DIMMs of comparable capacity.

DIMM error detection/correction technologies

Memory modules are inherently susceptible to memory errors. Each DRAM chip stores data in columns and rows of capacitors, or memory cells. The DIMM continuously refreshes the cells to preserve the data. The operating voltage of the memory device determines the level of the electrical charge. If an external event affects a capacitor's charge, the data may become incorrect. Such memory errors can cause applications and operating systems to crash, sometimes resulting in permanent data loss.

Memory errors are classified by the number of bits that are affected—single-bit or multi-bit—and the cause of error. A 64-bit wide data bus transports 64 bits at a time. These 64 bits constitute an ECC data word. An error in one bit of a data word is a single-bit error. An error in more than one bit of a data word is a multi-bit error.

Depending on the cause, engineers refer to memory errors as either hard or soft. Broken or defective pieces of hardware, such as DRAM defects, bad solder joints, and connector issues, cause hard errors so the device consistently returns incorrect results. For example, a memory cell may be stuck so that it always returns "0" bit, even when a "1" bit is written to it. Soft errors are more prevalent. They occur randomly when an electrical disturbance near a memory cell alters the charge on the capacitor. A soft error does not indicate a problem with a memory device because once the stored data is corrected the error does not recur.

The increasing possibility of memory errors

Two trends increase the likelihood of memory errors in servers:

- Expanding memory capacity
- Increasing storage density

Software vendors are developing increasingly complex, memory-intensive applications. This drives operating systems to address more memory, which causes manufacturers to expand memory capacity. Increased memory use increases the possibility of memory errors.

The storage density of the DRAM chips depends on the operating voltage of the memory system. As the size of memory cells decreases, both DRAM storage density and the memory-cell voltage sensitivity increase. Initially, industry-standard DIMMs operated at 5 volts. Because of improvements in DRAM storage density, operating voltage decreased first to 3.3 V, then 2.5 V, and then 1.8 V, allowing memory to run faster and consume less power. However, with increased memory storage and decreased operating voltage, a higher probability exists that an error may occur. Whenever a data bit is misinterpreted and goes uncorrected, the error can cause an application to crash. The only true protection from memory errors is to use some sort of memory detection or correction protocol. Some protocols only detect errors, while others can both detect and correct memory problems seamlessly.

Basic ECC memory

Parity checking detects only single-bit errors. It does not correct memory errors or detect multi-bit errors. HP introduced Error Correction Code (ECC) memory in 1993 and we continue to implement advanced ECC in

all HP ProLiant servers. ECC detects both single-bit and multi-bit errors in a 64-bit data word; it corrects single-bit errors.

ECC encodes information in a block of 8 bits to recover a single-bit error. When the DIMM writes data to memory, ECC uses a special algorithm to generate values called check bits. The algorithm adds the check bits together to calculate a checksum, which it stores with the data. When reading data from memory, the algorithm recalculates the checksum and compares it with the checksum of the written data. If the checksums are equal, then the data is valid and operation continues. If they differ, the data has an error, and the ECC memory logic isolates the error and reports it to the system. In the case of a single-bit error, the ECC memory logic can correct the error and output the corrected data so that the system continues to operate (Figure 6).

Figure 6. ECC logic locating and correcting a single-bit error
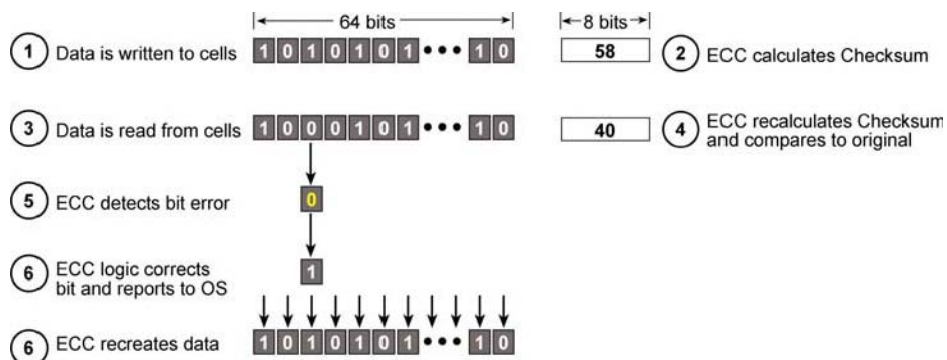

In addition to correcting single-bit errors, ECC detects, but does not correct, errors of two random bits and up to four bits within a single DRAM chip. ECC memory responds to these multi-bit errors by generating a Non-Maskable Interrupt (NMI) message that instructs the system to halt. ECC technology provided adequate protection for many applications. But the effectiveness of ECC protection decreases as memory capacity rises. These trends help to drive manufacturers to build more memory capacity in industry-standard servers:

• Operating system support for increasing amounts of memory
• Availability of low-cost, high-capacity memory modules
• Server virtualization

Advanced ECC memory

To improve memory protection beyond standard ECC, HP introduced Advanced ECC technology in 1996. HP and most other server manufacturers use this solution in industry-standard products. Advanced ECC can correct a multi-bit error that occurs within a DRAM chip and avoid a complete DRAM chip failure. In Advanced ECC with 4-bit memory devices, each chip contributes four bits of data to the data word. The four bits from each chip are distributed across four ECC devices (one bit per ECC device), so that an error in one chip could produce up to four separate single-bit errors. Figure 7 shows how one ECC device receives four data bits from four DRAM chips.

Figure 7. Advanced ECC
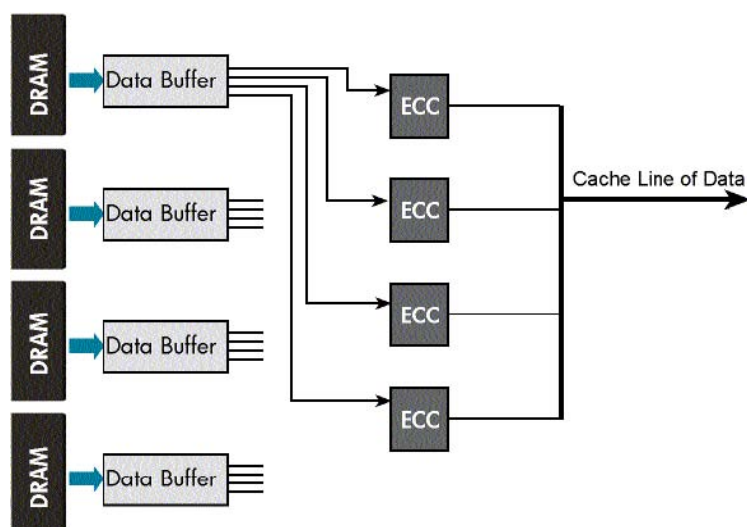
Since each ECC device can correct single-bit errors, Advanced ECC can correct a multi-bit error that occurs within one DRAM chip. As a result, Advanced ECC gives you protection from device failure (Table 1).

Table 1. Comparison of ECC and Advanced ECC error protection

Error Condition	ECC Outcome	Advanced ECC Outcome
Single-bit	Correct	Correct
Double-bit	Detect	Correct or detect
DRAM failure	Detect	Correct

Although Advanced ECC provides failure protection, it can reliably correct multi-bit errors only when they occur within a single DRAM chip. Advanced ECC does not provide failover capability. If there is a memory failure, you must shut down the system before replacing the memory. The latest generation of HP ProLiant servers offers several forms of Advanced Memory Protection that provide increased fault tolerance for applications requiring higher levels of availability.

Memory protection technologies

Advanced ECC corrects memory errors, but it does not give your system failover capability. Replacing a failed DIMM usually raises your operating costs when you take a server down for unscheduled maintenance, and the shutdown loses revenues for your business. Three available memory protection technologies, known as Memory Failure Recovery, give you failover and backup capability:

- Online Spare Memory mode
- Mirrored Memory mode
- Lockstep Memory mode

Online Spare Memory mode

In Online Spare Memory mode, you can designate a populated channel (or branch) as the spare, making it unavailable for normal use as system memory. If a DIMM in the system channel exceeds a threshold rate of

correctable memory errors, the channel goes off line and the system copies data to the spare channel. This prevents data corruption, a server crash, or both, and you can replace the defective DIMM at your convenience during a scheduled shutdown. This also maintains server availability and memory reliability without service intervention or server interruption. Online Spare Memory technology reduces the chance of an uncorrectable error bringing down the system; however, it does not fully protect your system against uncorrectable memory errors.

NOTE

Online Spare Memory mode can run on some systems with only one memory channel populated. But a single-channel memory configuration requires dual-rank DIMMs. For more information, refer to the server's user guide.

In a system with three channels per memory controller, two channels operate normally and the third channel is the spare. Online Spare Memory mode does not require operating system support or special software beyond the system BIOS. However, to support messaging and logging at the console along with messages in HP Systems Insight Manager, the operating system must have system management and agent support for Advanced Memory Protection. Implementing Online Spare Memory mode over Advanced ECC requires extra DIMMs for the spare memory channel and reduces the system's memory capacity.

Mirrored Memory mode

Mirrored Memory mode is a fault-tolerant memory option that gives you a higher level of security than Online Spare mode. Mirrored Memory mode provides full protection against single-bit and multi-bit errors.

With Mirrored Memory mode, the memory subsystem writes identical data to two channels simultaneously. If a memory read from one of the channel returns incorrect data due to an uncorrectable memory error, the system automatically retrieves the data from the other channel. A transient or soft error in one channel does not affect mirroring, and operation continues unless there is a simultaneous error in exactly the same location on a DIMM and its mirrored DIMM. Mirrored Memory mode reduces the amount of memory available to the operating system by 50% since only one of the two populated channels provides data.

Lockstep Memory mode

Lockstep Memory mode uses two memory channels at a time and offers you an even higher level of protection. In lockstep mode, two channels operate as a single channel—each write and read operation moves a data word two channels wide. Both channels split the cache line to provide 2x 8-bit error detection and 8-bit error correction within a single DRAM. In three-channel memory systems, the third channel is unused and left unpopulated. The Lockstep Memory mode is the most reliable, but it reduces the total system memory capacity by one-third in most systems.

Memory protection mode summary

The Online Spare Memory, Mirrored Memory, and Lockstep Memory modes enhance reliability and availability. A faulty DIMM does not have to result in unscheduled downtime. The server can remain up and running until it is convenient to take the equipment off line. Unfortunately, memory protection modes sacrifice performance. The reduction in available channels limits bandwidth and reduces a system's capacity. Nevertheless, server availability is the highest priority for enterprise environments.

You can configure the system for the desired memory protection mode through the RBSU. Online Spare, Mirrored, and Lockstep memory protection modes are mutually exclusive, meaning that only one mode operates at a time. You must use DIMMs that are identical in type, size, and rank.

Advanced memory technologies

Advanced memory technologies fill the growing performance gap between memory and processor despite the gains made possible by SDRAM. The industry has developed several technologies to boost system performance using the latest high-speed processors (Figure 8).

Figure 8. Peak bandwidth comparison of SDRAM and advanced SDRAM technologies
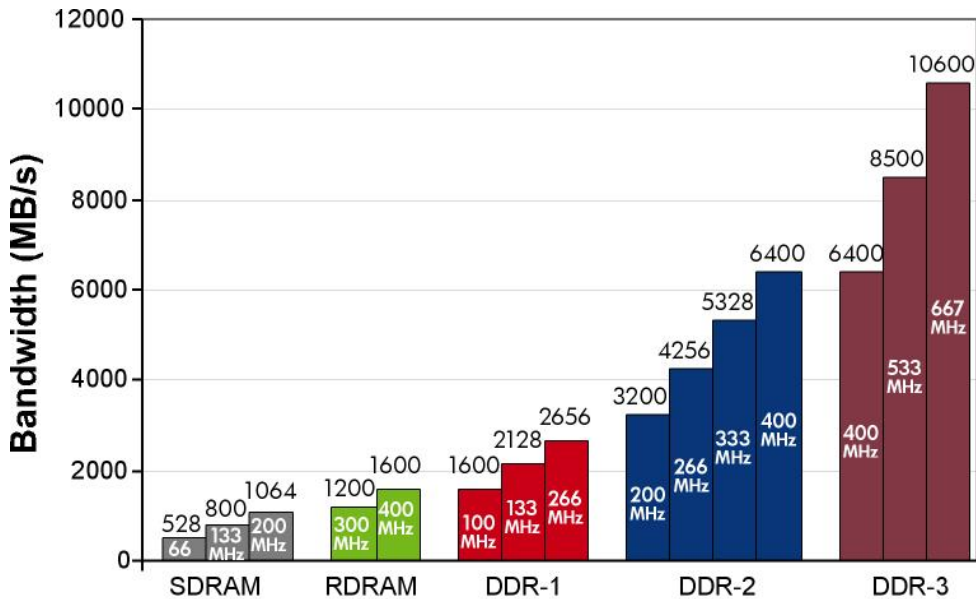
Rambus DRAM

Rambus DRAM (RDRAM) is a memory technology introduced in the late 1990s to achieve faster memory throughput than SDRAM. Unlike SDRAM, Rambus used a high-speed serial connection to move data from memory modules to the system. The Rambus design consisted of three key elements: RDRAMs, Rambus ASICs, and a Rambus Channel interconnect. Unlike the later FB-DIMM standard, the Rambus system used a multi-drop serial bus to communicate with each RDRAM chip rather than using an advanced memory buffer (AMB).

RDRAM delivered better memory bandwidth than SDRAM, but it required a different memory bus design and significantly more expensive RDRAM chips. Double Data Rate (DDR) SDRAM, with a design similar to SDRAM but having higher bandwidth and lower costs, replaced RDRAM in server designs.

Double Data Rate SDRAM technologies

Double Data Rate (DDR) SDRAM technology doubles the SDRAM data transfer rate without increasing the frequency, or speed, of the memory clock. This section describes three generations of DDR SDRAM technology.

DDR1

To develop the first generation of DDR SDRAM (DDR1), designers made enhancements to the SDRAM core to increase the data rate. These enhancements include prefetching, double transition clocking, strobe-based

data bus, and Stub-Series Terminated Logic_2 (SSTL_2) low-voltage signaling. At 400 MHz, DDR increases memory bandwidth to 3.2 GB/s—400% more than original SDRAM.

Prefetching

SDRAM transfers one bit per clock cycle from the memory cell array to the I/O buffer or data queue (DQ). The I/O buffer releases one bit to the bus per pin and clock cycle (on the rising edge of the clock signal). To double the data rate, DDR SDRAM uses a technique called prefetching to transfer two bits from the memory cell array to the I/O buffer in two separate pipelines. Then the I/O buffer releases the bits in the order of the queue on the same output line. This is a 2n-prefetch architecture because the DRAM fetches two data bits from the memory cell array before releasing them to the bus in a time-multiplexed manner.

Double transition clocking

Standard DRAM transfers one data bit to the bus on the rising edge of the bus clock signal, while DDR1 uses both the rising and falling edges of the clock to trigger the data transfer to the bus. This technique, known as double transition clocking, delivers twice the bandwidth of SDRAM without increasing the clock frequency. DDR1 has theoretical peak data transfer rates of 1.6 and 2.1 GB/s at clock frequencies of 100 MHz and 133 MHz, respectively.

Stub-Series Terminated Logic_2 (SSTL_2) low-voltage signaling technology

The signaling technology is another difference between SDRAM and DDR1. Instead of using a 3.3-V operating voltage of SDRAM, DDR1 uses the 2.5-V signaling specification SSTL_2. This low-voltage signaling results in lower power consumption and improved heat dissipation.

Strobe-based data bus

SSTL_2 signaling lets DDR1 run faster than traditional SDRAM. DDR1 also uses a delay-locked loop (one for every 16 outputs) to provide a data strobe signal as data becomes valid on the SDRAM pins. The memory controller uses the data strobe signal to locate data more accurately and resynchronize incoming data from different DIMMs.

DDR1 operates at transfer rates of 400 Mb/s, or 3.2 GB/s. Although the data bus is capable of running at these speeds, the command bus cannot. The use of strobes alleviated tight system timing requirements on the data bus. But the command bus does not use a strobe and must still meet setup times to a synchronous clock, so at a data rate of 400 Mb/s, the command bus must operate at 200 MHz.

DDR1 DIMMs

DDR1 DIMMs require 184 pins instead of the 168 pins used by standard SDRAM DIMMs. DDR1 is versatile enough for both desktop PCs and servers. To vary the cost of DDR1 DIMMs for these different markets, memory manufacturers offer unbuffered and registered versions. Unbuffered DDR1 DIMMs place the load of all the DDR modules on the system memory bus. Use Unbuffered DIMMS in systems that do not require high memory capacity. On the other hand, registered DDR1 DIMMs place only one load per DIMM on the memory bus, regardless of how many SDRAM devices are on the module. They are best suited for servers with high memory capacities.

DDR2

DDR2 is the second generation of DDR SDRAM technology. It provides data rates of up to 6.4 GB/s, lower power consumption, and improvements in packaging. At 400 MHz and 800 Mb/s, DDR2 increases memory bandwidth to 6.4 GB/s—800% more than original SDRAM. DDR2 SDRAM achieves higher performance level and lower power consumption through faster clocks, 1.8-V operation and signaling, along with a simplified command set. The 240-pin connector on DDR2 accommodates differential strobe signals (Figure 9).

Figure 9. DDR2 DIMM with 240-pin interface
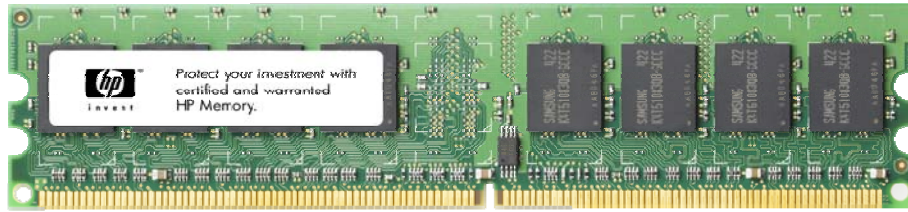
DDR3

DDR3, the third-generation of DDR SDRAM technology, makes further improvements in bandwidth and power consumption. DDR3 manufacturers began fabrication using with 90 nm technologies. With increasing production volumes, they are moving toward 70 nm technology. DDR3 operates at clock rates from 400 MHz to 1066 MHz with theoretical peak bandwidths ranging from 6.40 GB/s to 17 GB/s. DDR3 DIMMs can reduce power consumption by up to 30% compared to DDR2 DIMMs operating at the same speed. DDR3 DIMMs use the same 240-pin connector as DDR2 DIMMs, but the notch key is in a different position (Figure 10).

Figure 10. DDR3 DIMM with 240-pin interface

To increase performance and reduce power consumption, designers made several key enhancements:

- An 8-bit prefetch buffer stores more data than the 4-bit buffer for DDR2.
- Fly-by topology for the commands, addresses, control signals, and clocks improves signal integrity by reducing the number of stubs and their length. This feature requires the controller to support "write leveling" on DDR3 DIMMs.
- 1.5-V signaling reduces power consumption from the 1.8- V signaling for DDR2. DDR3 also supports low voltage DIMMs that use 1.35-V signaling and consume 10% to 15% less power than standard DDR3 DIMMs.
- A thermal sensor integrated on the DIMM module signals the chipset to throttle memory traffic if DIMM temperature exceeds a programmable critical trip point.

Module naming convention and peak bandwidth

Table 2 summarizes the various types of DDR1, DDR2, and DDR3 SDRAM as well as their associated naming conventions. Originally, the effective clock rate of the data transfer determined the module naming convention for DDR-SDRAM. For example, a PC200 designation stands for DDR SDRAM that operates at 100 MHz, while PC266 stands for 133 MHz. Dueto confusion over the Rambus naming convention, the

industry based the DDR-SDRAM naming convention on the actual peak data transfer rate in MB/s. For example, PC266 is equivalent to PC2100 (64 bit * 2 * 133 MHz = 2.1 GB/s or 2100 MB/s).

DDR3 SDRAM DIMM manufacturers produce two types of DIMMs: Unbuffered DIMMs (UDIMM) and Registered DIMMs (RDIMM). UDIMMs are the most basic type of memory module and offer a lower latency and low power consumption but are limited in capacity. An E suffix in the manufacturer's module name identifies Unbuffered DIMMs with ECC.

RDIMMs offer larger capacities than UDIMMs and include address parity protection. An R suffix in the module manufacturer's name identifies Registered DIMMs.

**Table 2**. Summary of DDR SDRAM technologies

Type	Component naming convention	Module naming convention	Bus speed	Peak bandwidth
DDR1	DDR200	PC1600	100 MHz	1.6 GB/s
	DDR266	PC2100	133 MHz	2.1 GB/s
	DDR333	PC2700	166 MHz	2.7 GB/s
	DDR400	PC3200	200 MHz	3.2 GB/s
DDR2	DDR2-400	PC2-3200R	200 MHz	3.2 GB/s
	DDR2-533	PC2-4300	266 MHz	4.3 GB/s
	DDR2-667	PC2-5300	333 MHz	5.3 GB/s
	DDR2-800	PC2-6400	400 MHz	6.4 GB/s
DDR3	DDR3-800	PC3-6400	400 MHz	6.4 GB/s
	DDR3-1066	PC3-8500	533 MHz	8.5 GB/s
	DDR3-1333	PC3-10600	667 MHz	10.6 GB/s
	DDR3-1600	PC3-12800	800 MHz	12.8 GB/s

Fully-buffered DIMMs

Traditional DIMM architectures use a stub-bus topology with parallel branches (stubs) that connect to a shared memory bus (Figure 11). Each DIMM connects to the data bus using a set of pin connectors. For the electrical signals from the memory controller to reach the DIMM bus-pin connections at the same time, all the traces must be the same length. This can result in circuitous traces on the motherboard between the memory controller and memory slots. Both the latency resulting from complex routing of traces and signal degradation at the bus-pin connections cause the error rate to increase as the bus speed increases.

Figure 11. Stub-bus topology
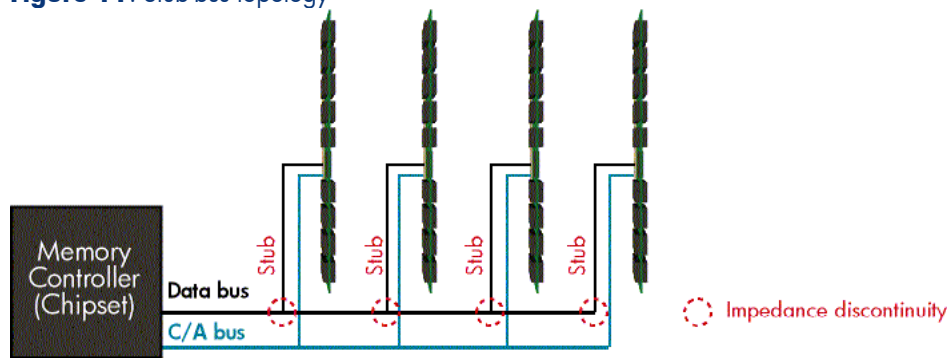
Each stub-bus connection creates resistance that can degrade signal integrity. In addition, each DIMM creates an electrical load on the bus. Adding more DIMMs increases the electrical load. As the bus speed increases, these factors decrease the number of supportable DIMMs per channel. For example, Figure 12 shows the number of loads supported per channel at data rates ranging from PC 100 to DDR3 1600. Note that the number of supported loads drops from eight to two as data rates increase to DDR2 800.

Figure 12. Maximum number of loads per channel based on DRAM data rate
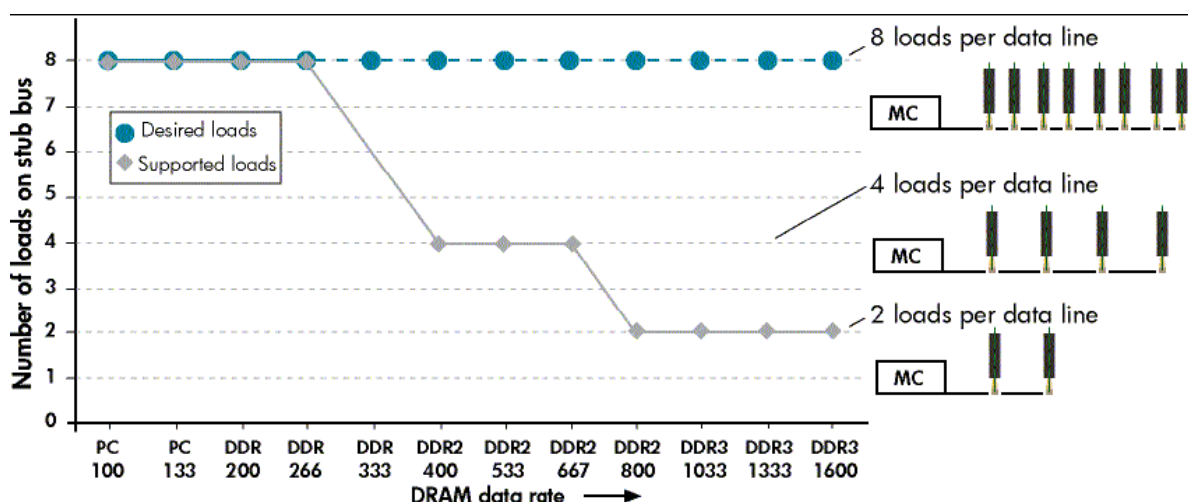
Due to increased cost and board complexity, increasing the number of channels to compensate for the drop in capacity per channel was not a viable option. System designers had two choices: limit memory capacity so that fewer errors occur at higher speeds, or use slower bus speeds and increase the DRAM density. For future generations of high-performance servers, neither option was acceptable.

New server designs require improved memory architecture to achieve higher memory bandwidth and capacity. The JEDEC developed the Fully-Buffered DIMM (FB-DIMM) specification, a serial interface that eliminates the parallel stub-bus topology and allows higher memory bandwidth while maintaining or increasing memory capacity.

The FB-DIMM architecture has serial links between the memory controller and the FB-DIMMs connected in a daisy chain configuration (Figure 13). Relative to the memory controller, there are ten outbound links and fourteen inbound links, also known as southbound and northbound links, respectively. These serial links

connect the memory controller to an Advanced Memory Buffer (AMB) chip that resides on each FB-DIMM, creating a point-to-point architecture. The outbound links transmit commands and write data to the FB-DIMMs while the inbound links transmit read data back to the memory controller.

The clock signal travels over a different set of pins. In addition to communicating over the outbound lanes, the memory controller communicates configuration information with each AMB over the System Management bus or SMBus. The AMB is an intelligent chip that manages serial communication with the memory controller and parallel communication with local DRAM devices. Each AMB receives signals for address, write data, and command information through the outbound links and re-transmits the signal to the next FB-DIMM on the channel. Each AMB decodes the command data and ignores the commands targeted for a different DIMM. The targeted AMB performs a read or write operation to local DRAM devices through a parallel interface. In a read operation, the AMB serializes data from the DRAM devices and transmits it to the memory controller through the inbound links.

Figure 13. Serial communication between daisy-chained FB-DIMMs on a single channel

Option
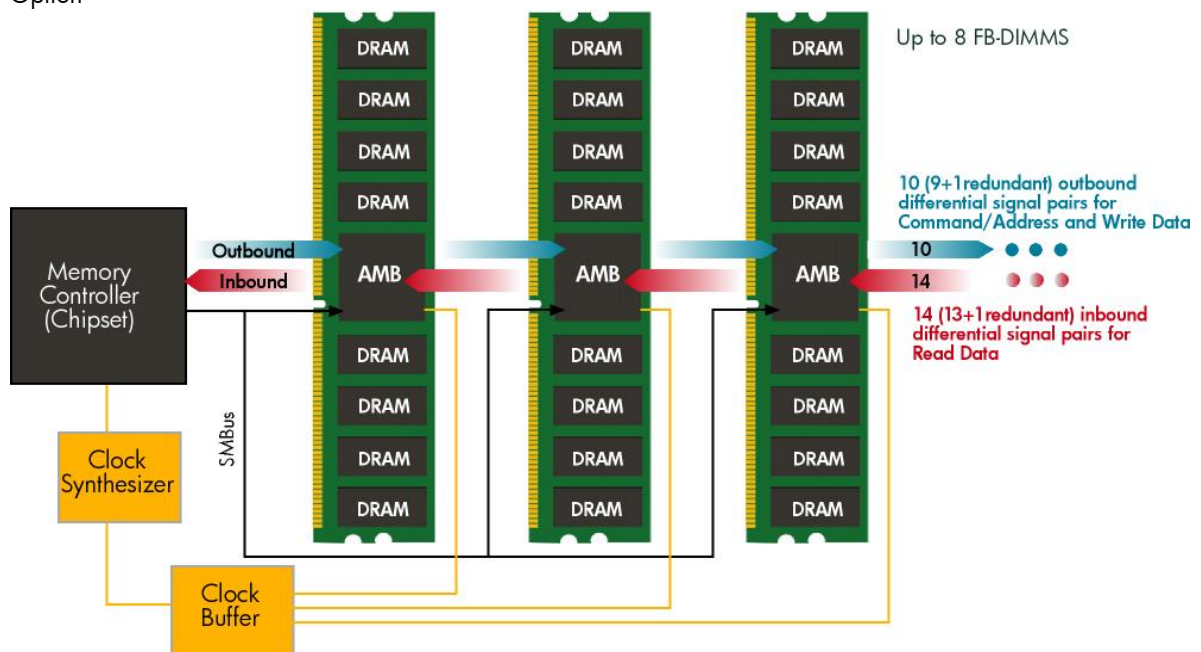
NOTE: AMD Opteron™and Intel® Xeon® E55xx/X55xx CPU designs include the memory controller and clock functions integrated into processor module.

When using DDR2-667 DRAM on the FB-DIMM, the peak theoretical throughput of the inbound links is 5.4 GB/s. The peak theoretical throughput of the outbound links is half that, approximately 2.6 GB/s.

With DDR2 memory technology, manufacturers implemented FB-DIMMs as well as Unbuffered and Registered DIMMs. In spite of its advantages,however, FB-DIMM memory also costs more, uses more power, and has increased latency. As a result, the industry has not implemented FB-DIMMs for DDR3. Instead, engineers have increased server memory capacities by designing systems with Non-Uniform Memory Access (NUMA) architectures, which feature up to four memory channels for each CPU.

Importance of using HP-certified memory modules in ProLiant servers

There are several reasons why you should use only HP memory option kits to replace or add memory in ProLiant servers.

First, not all DIMMs are created equal. They can vary greatly in quality and reliability. In the highly competitive memory market, some third-party memory resellers forego the level of qualification and testing needed for servers because it adds to the price of DIMMs. We use proprietary diagnostic tools and specialized server memory diagnostic tests that exceed industry standards to ensure the highest level of performance and availability for HP ProLiant servers. The cost related to system downtime, data loss, and reduced productivity caused by lower quality memory is far greater than the price difference between HP-certified memory and third party DIMMs.

Second, HP offers three levels of AMP that go beyond error correction to increase fault tolerance. These AMP technologies—Online Spare Memory, Mirrored Memory, and Lockstep Memory—are optimized for each server series. For ProLiant servers with AMP, the AMP mode you select determines the DIMM configuration. HP memory option kits ensure correct configuration for each AMP mode.

Third, using HP memory option kits prevents improper mixing of single-, dual-, and quad-rank DIMMs. Although DIMMs of different ranks may have the same capacity, they differ in the way the memory controller accesses them. To ensure that the server boots properly and operates reliably, each channel's DIMM load should be consistent. HP memory option kits precisely match the capabilities and requirements of the designated ProLiant server.

To find the correct HP memory option kit, consult the server's user guide or the product QuickSpecs available at www.hp.com.

Conclusion

The constant need to boost memory performance for increasingly powerful system processors drives the development of advanced memory technologies. At HP, we continue to evaluate relevant memory technologies to give our customers products with the most reliable, best performing memory at the lowest possible cost.

For more information

For additional information, refer to the resources listed below.

Resource description	Web address
JEDEC Web site	http://www.jedec.org
HP ProLiant Server Memory web page	http://h18004.www1.hp.com/products/servers/options/memory-description.html
DDR3 memory technology, technology brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c02126499/c02126499.pdf

Call to action

Send comments about this paper to TechCom@HP.com